Applied Study of Medical Image Segmentation Based on Convolutional Neural Network

Shuai Ding[1], Zhiyang Huang[2, *]

[1]Fuzhou NO.3 Middle School, Fuzhou, China

[2]Nantou Middle School, Shenzhen,China

*Corresponding author: 1603011425@stu.hrbust.edu.cn

Abstract. Medical image segmentation is one of the hot topics in today's research, and researchers have found that medical image segmentation has important application prospects and potentials in the diagnosis and treatment of diseases, however, there are still some challenges and problems in the practical application. Therefore, the research topic of this paper is the research of medical image segmentation applications based on convolutional neural networks. The research methodology of this paper is to analyze the application of existing deep learning algorithms in image segmentation. It is found that the medical image segmentation method based on the convolutional neural network can effectively improve the accuracy and efficiency of segmentation results. The reliability of segmentation can be further improved by introducing medical experts' knowledge and labeled data. In summary, the research in this paper shows that the medical image segmentation method based on convolutional neural networks has the potential and advantages to improve the accuracy and efficiency of segmentation.

Keywords: Deep learning; image segmentation; convolutional neural networks.

1. Introduction

Medical image segmentation is one of the important tasks in the field of medical imaging, aiming to segment structures and tissues in medical images into different regions for applications such as disease diagnosis, surgical planning, and treatment monitoring. Traditional medical image segmentation methods are usually based on hand-designed feature extraction and classification algorithms, but these methods have some limitations when dealing with complex medical images, such as the need for a large number of human interventions, high requirements on image quality, and slow processing speed.

In recent years, with the development of artificial intelligence technology, especially the rise of convolutional neural networks, significant progress has been made in the field of medical image segmentation. CNN is a deep learning model that can automatically learn features in medical images and achieve efficient segmentation of images through multi-layer convolution and pooling operations.

Dou et al. applied a strongly supervised method to segment the volume of three-dimensional liver CT. This is achieved by using deconvolution layers to the middle and features and applying SoftMax layers to encrypt the classification output [1]. Their results not only demonstrate good convergence but also reduce training and validation errors. Various relationships among chest X-ray images, including interpersonal, age-related, gender-related, and view-related aspects, have been utilized for graph construction. The inductive learning problem was addressed using GraphSAGE [2], while the multi-relation fusion issue was tackled using relational GCN [3]. Du et al. integrated CNN with GCN, adopting the GAT algorithm to discriminate lesions in ROI features extracted via CNN [4]. This approach effectively simulated radiologists' magnification of lesion ROIs, culminating in the use of GAT for breast cancer detection in X-ray images.

In another study on breast cancer detection, Ye et al. divided images into different ROI blocks [5], initially utilizing U-Net for tumor segmentation in ROIs, followed by the use of GCN to capture the topological structure of ROI images [6]. Subsequently, a fully connected network was employed for feature classification, achieving breast cancer detection in ROI areas.

This study aims to deepen the understanding of the application of deep learning in the medical field and to provide a scientific basis and methods for improving medical diagnosis and treatment methods, promoting medical research, and improving public health.

2. Deep Learning Algorithm Application

Strong supervision is a machine learning method that requires the use of datasets with clear labels during its training process. In strongly supervised learning, the model learns the mapping relationship between input data and associated labels to make accurate predictions on new data that has not been seen before. This method applies to many tasks, including classification, regression, and object detection.

Weak supervision is a concept relative to strong supervision, which refers to the use of only partial, incomplete, or imprecise label information during the training process. In weakly supervised learning, labels may be imprecise, noisy, or only involve a portion of the data. The goal of weakly supervised learning is to train models using limited label information to perform well in dealing with more challenging tasks.

Learning falls between strong supervision and weak supervision. In learning, the training process utilizes both labeled and unlabeled data simultaneously. Usually, labeled data is scarce, while unlabeled data is relatively abundant. The model aims to utilize the information from unlabeled data during the learning process to improve the learning effectiveness of labeled data. learning is commonly used to deal with situations where data annotation costs are high, or it is difficult to obtain large amounts of labeled data.

In the field of deep learning, these learning methods have a wide range of applications. Strong supervision is used to train deep neural networks, while weak supervision and learning methods fully consider the situation where label data is insufficient or difficult to obtain in real scenarios, to improve the generalization performance of the model. These methods play important roles in tasks such as image classification, object detection, and semantic segmentation.

The core idea of strong supervision is to directly supervise the hidden layer and propagate it to the lower layer, rather than just at the output layer. This idea achieves non-medical purposes by adding accompanying objective functions in the hidden layer. Similarly, in Google Net, two hidden layers of the 22-layer network were supervised.

In a similar method, three classifiers were injected to classify intermediate output features from the contracted part of the class U-Net structure. The classification output is used as a regulator during the training phase. The multi-level contextual information in the network helps improve localization and discrimination abilities. In addition, the auxiliary classifier enhances the backpropagation flow of gradients during the training phase.

Through supervised learning, computer systems are able to learn from a large number of images with labeled information. The computer performs accurate object detection in unlabeled new images through deep learning. The successful application of this method involves several key steps and techniques. In general, the application of supervised learning in image detection has achieved remarkable results.

Through reasonable selection of data and training strategies, the target detection model can be trained to perform well in various complex scenarios. However, with the increasing complexity of the problem, the research and improvement of supervised learning in image detection is still a continuous development field

3. Results and Discussion

Many of the image segmentation uses strong supervision, but strong supervision requires explicit labels to guide the training of the network, it will consume a lot of time and money, so we can try weak supervision to accomplish the task of image segmentation. Weakly supervised image segmentation methods provide a cost-effective way to solve the difficult and costly problem of medical image analysis. This approach typically combines a small amount of pixel-level data with a large amount of unlabelled data for training.

By making full use of unlabelled data, the model learns a more generalized feature representation, which improves its adaptability to new data. This improved generalization ability is particularly critical for medical image segmentation tasks, where there are large individual differences in patient physiology and lesions. Weak supervision requires only a small amount of pixel-level data and a large amount of unlabelled data for training, making it a cost-effective approach.

In addition, the ultimate goal of our task is to cure the patient, so it is possible to combine our images with the corresponding textual information to make a multimodal task, which can provide a corresponding solution based on the analysis of the current lesion area and its category.

The application of semi-supervised learning in the field of medical image segmentation is obviously a very effective strategy, especially in the case of limited data. Medical images usually require professional domain knowledge for accurate data, which is costly and difficult to obtain data. Semi-supervised learning has the advantage of being able to make fuller use of unlabelled data, through which the model can be optimized and constrained to improve model performance.

In semi-supervised learning, data is used to supervise the learning of the model, while unlabelled data is used for model constraints and optimization. By making full use of the information from unlabelled data, semi-supervised learning can improve model performance with limited data and can balance the learning effects of strong and weak supervision.

Convolutional neural networks can be used for target detection in medical images, such as detecting lesions. By sliding a convolutional window in an image, CNNs can locate and identify target regions in an image, providing more accurate information about the location and size of lesions. This method can provide doctors with more comprehensive information, including not only the location and shape of the lesion but also textual descriptions related to the image.

This multimodal information can help doctors gain a deeper understanding of the patient's condition, providing support for developing more personalized and effective treatment plans. Multimodal tasks also help the model to interpret clinical decisions and improve its usability in practical medical environments.

In summary, the successful application of Convolutional Neural Networks (CNNs) in the detection of targets within medical imaging has marked a significant advancement in the field of medical image analysis. Through convolutional operations, CNNs effectively capture local features in images, enabling precise detection of tumors, lesions, and other targets.

Utilizing a sliding convolutional window approach, CNNs can locate and recognize target areas in images, providing more accurate information on the location and size of lesions, thus offering reliable diagnostic support to physicians. The strength of CNNs lies in their ability to automatically learn feature representations, eliminating the need for manually designed features, and making them well-suited for complex and varied medical imaging data.

In practice, it is imperative to maintain a high level of attention to data privacy and ethical issues. Medical imaging involves patient privacy, necessitating effective measures for privacy protection in image processing and model training. Additionally, the interpretability of models is a key concern in medical applications.

Physicians need to understand the decision-making process of models to better apply the output results in clinical practice. The quality and variety of data are crucial to the success of supervised learning in image detection. The training data should cover a variety of scenarios and target shapes to ensure good generalization of the model. The accuracy of labeled data is crucial for training high-performance detection models. The use of accurate annotation information can effectively guide the model to learn the features and location of the target.

For image detection, the choice of loss function is also very important. Common loss functions include cross-entropy loss, smooth L1 loss, etc. These loss functions can effectively guide the model to learn the classification and location information of the object. The balance of positive and negative samples and the focus on hard-to-detect targets are also important factors to consider in the design of loss functions.

In the training process, the application of data enhancement technology can increase the diversity of training samples. Common data enhancement operations include random rotation, cropping, scaling, and flipping. These actions help the model better adapt to different scenarios and conditions.

4. Conclusion

Convolutional Neural Networks have been extensively applied in medical image segmentation tasks, such as tumor localization and organ segmentation. Owing to their effective capture of local features in images, CNNs are highly effective for accurately locating and segmenting structures within medical images. The successful application of CNNs in the field of medical image segmentation has provided powerful tools for clinical medicine and medical research. Their high-precision segmentation results offer more accurate diagnostic information to physicians, advancing medical image processing technology.

In addition to convolutional neural networks, fully convolutional networks have DeepLab series, PSPNet, region-based convolutional neural networks, and HRNet for image segmentation with applications. Capturing contextual information at different scales through the pyramid pooling module helps in better understanding the semantic information of the image. Region-based convolutional neural networks, combine target detection and semantic segmentation, allowing for simultaneous detection and segmentation of objects in an image.

Future research will investigate that these algorithms are usually trained by training on large-scale datasets, such as Cityscapes to learn representations for extracting semantic information from images. The selection of appropriate algorithms usually depends on the application scenario, computational resources, and performance requirements.

However, there remain challenges to be addressed, including issues related to data privacy, model interpretability, and the standardization of uncertainty estimation. Future research directions will focus on addressing these challenges while further enhancing the practicality and clinical applicability of the models. Continuing exploration of new methods and technologies in the field of medical image segmentation with convolutional neural networks remains a necessity for future research.

Authors Contribution

All the authors contributed equally and their names were listed in alphabetical order.

References

[1] Tang A , Wartellebladou C , Gilbert G ,et al.Response to Comment on Tang et al. Effects of Insulin Glargine and Liraglutide Therapy on Liver Fat as Measured by Magnetic Resonance in Patients With Type 2 Diabetes: A Randomized Trial. 2015.

[2] Julie,Edge,Carlo,et al.An alternative sensor-based method for glucose monitoring in children and young people with diabetes.[J]. Archives of disease in childhood, 2017.

[3] Schlichtkrull M , Kipf T N , Bloem P ,et al.Modeling Relational Data with Graph Convolutional Networks[J]. 2017.

[4] Zheng A , Wang H, Wang J,et al.Diverse features discovery transformer for pedestrian attribute recognition[J].Engineering Applications of Artificial Intelligence, 2023.

[5] Hou Y E, Li W G, Rong A Q, et al.Tracking algorithm of block sparse representation with background information[J].Journal of South China University of Technology(Natural Science Edition), 2013, 41(8):21-27.

[6] Diamond K M, Avants B B, Maga A M . Machine learning-based segmentation and landmarking of 2D fish images[J]. 2021.